



# Etude préliminaire à un modèle de prévision à court terme de l'activité d'un transporteur sous température dirigée

Wilfried Despagne

## ► To cite this version:

Wilfried Despagne. Etude préliminaire à un modèle de prévision à court terme de l'activité d'un transporteur sous température dirigée. 2008. hal-00238468

**HAL Id: hal-00238468**

**<https://hal.science/hal-00238468>**

Preprint submitted on 4 Feb 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Etude préliminaire à un modèle de prévision à court terme de l'activité d'un transporteur sous température dirigée

Wilfried Despagne\*

\*Laboratoire SABRES,  
Agrostar (le pôle systèmes d'information du Groupe STEF-TFE)  
Wilfried.Despagne@stef-tfe.com

**Résumé.** Cet article décrit une problématique de recherche opérationnelle. Un transporteur sous température dirigée cherche à optimiser la planification de ses ressources humaines et matérielles à travers la prévision à très court terme de son activité. Le challenge réside dans le fait de trouver un modèle de prévision unique s'adaptant, sans intervention humaine, aux spécificités des 57 agences du transporteur. La matière première est l'information récoltée par le transporteur depuis plus de six ans. Les outils sont des algorithmes mathématiques utilisés pour la prévision des séries temporelles. Le travail décrit ici vise à combiner ces outils pour qu'ils extraient le maximum d'information déterministe capable d'être anticipée. L'introduction pose la problématique et son contexte économique. Elle est suivie d'un descriptif des procédures utilisées et d'un argumentaire pour défendre leur choix. Les solutions informatiques adoptées sont inventoriées. Enfin, la conclusion renvoie à des pistes d'études.

## 1 Introduction

La présentation suivante est le préliminaire d'un travail de recherche dans le cadre d'une thèse Cifre. Les structures dans lesquelles ce travail est mené, sont le pôle systèmes d'information du groupe STEF-TFE appelé Agrostar et le laboratoire SABRES de l'Université Bretagne Sud. Nous présentons un état des lieux de la prévision d'activité chez un transporteur sous température dirigée. Cette démarche a permis d'acquérir un point de vue formalisé du problème et une automatisation des procédures.

Le transport sous température dirigée est l'activité qui consiste à transporter de la marchandise soumise à une température définie entre  $-25^{\circ}\text{C}$  et  $+15^{\circ}\text{C}$ . Les marchandises sont essentiellement des denrées alimentaires périssables, les produits carnés, les produits de la mer, les fruits et légumes, les produits laitiers, les surgelés, mais aussi des plantes ou médicaments. Elles ont en commun d'être soumises à un cahier des charges très strict qui définit la « chaîne du froid ». La chaîne du froid est le processus qui permet de maintenir un produit à basse température. Le froid ralentit la propagation des microorganismes. La loi impose des règles en cette matière. Les arrêtés du 9 mai 1955 (réglementant l'hygiène des aliments remis directement au consommateur) et du 20 juillet 1998 (fixant les conditions techniques et

## Prévisions d'activité à des fins opérationnelles

hygiéniques applicables au transport des aliments) soumettent les industriels à une obligation de résultat. À la contrainte du maintien de la température liée à la chaîne du froid, s'ajoutent celles liées au métier du transporteur. Là encore, le législateur réglemente la durée du travail d'un chauffeur routier et l'autorisation de circulation. Ce ne sont ici que des exemples de lois parmi d'autres permettant de percevoir l'ampleur des contraintes législatives en matière de transport et plus encore de transport frigorifique. Sans une maîtrise de gestion des ressources matérielles et humaines, ces contraintes font déborder les coûts de fonctionnement.

Le domaine de l'alimentaire met en relation producteurs, industriels, distributeurs et transporteurs. Ils forment une chaîne appelée « Supply Chain » (chaîne logistique, cf Ayadi (2005)). La contrainte consistant à garder la fraîcheur des produits entre leur lieu de fabrication et celui de distribution, impose aux différents maillons de la chaîne de travailler en flux tendus. La Date Limite de Consommation (DLC) est de quelques jours pour des découpes de volaille ou de 21 jours pour des yaourts. Elle gouverne la « Supply Chain » des produits frais et imprime un rythme rapide à la chaîne de distribution. Dans son mémoire, Charlotte Terrolle (2004) souligne que les industriels répondent aux commandes quotidiennes des GMS (Grandes et Moyennes Surfaces) pour approvisionner leurs entrepôts. Les quantités de ravitaillements fluctuent quotidiennement selon les sorties de caisse des GMS. Dans le cas de Carrefour, tous les soirs, toutes les données de vente en magasin sont centralisées à l'entrepôt, les commandes sont expédiées dès le lendemain. Ce réapprovisionnement automatique représente : 80% des ventes pour l'épicerie et 98% des produits frais. Pour soutenir ce rythme et alimenter les linéaires sans pour autant avoir de stocks, les acteurs doivent prévoir leurs ventes. La logique voudrait que les GMS partagent leurs prévisions de ventes, aussi bien à un rythme mensuel que quotidien, avec les industriels de l'agroalimentaire et que ces derniers les relient aux transporteurs. Mais la loi de la concurrence ne permet pas cette pratique. Les GMS craignent de donner trop d'informations aux industriels qui, se voyant dans une position de force, pourraient en profiter pour augmenter leurs prix. Ainsi, les acteurs de la chaîne logistique établissent des prévisions chacun de leur côté.

Dans le cadre de ce travail, nous proposons un système de prévision des ventes du transporteur sous température dirigée TFE. Il est la branche transport du groupe STEF-TFE, leader français du secteur de la logistique du froid. TFE dispose d'un réseau 57 agences en Europe. Une agence comporte une plate-forme (ou quai) sur laquelle est réceptionnée la marchandise. Ces quais permettent au transporteur de préparer les commandes, les trier et les étiqueter pour les dépêcher à travers l'Hexagone en moins de 48h. Dans le jargon métier c'est un transport en A pour B. Grossièrement, il existe une agence par région. D'une part, elle est en charge d'enlever la marchandise chez ses clients régionaux pour l'injecter dans le réseau qui s'occupe de la livrer à bon port. De même, elle reçoit de la marchandise du réseau et la livre aux points de vente de sa région.

Comme les autres maillons de la « Supply Chain » alimentaire, TFE travaille suivant la méthode JAT (Juste A Temps). Dans plus de 90% des cas, l'agence reçoit du client expéditeur ses ordres de transports moins de 3h avant l'enlèvement de la marchandise. Pour résumer en quelques mots la problématique, les ordres de transports tombent par fax ou EDI (Échange de Données Informatisé) à 7h du matin pour un départ entre 8h et 12H en fonction de la destination. La livraison doit se faire entre 18h et 22h le soir sur toute la France. À cause des opérations spéciales (promotions), il arrive que les quantités à transporter varient de 1 à 10 d'un jour à l'autre. Dans ces cas, comment faire face si les clients n'en communiquent pas la

période ?

Le transporteur doit donc s'adapter aux exigences de ses clients. Pour l'aider, il souhaite mettre en place un système de prévision pour l'aide à la gestion des ressources. Le système doit permettre d'anticiper le poids des marchandises à transporter et le nombre de lettres de transports à couvrir. Ces deux informations, prévues à un horizon de 15 jours et à périodicité journalière, lui permettent d'anticiper l'effectif humain et matériel sur le quai ainsi que le nombre de semi-remorques à mettre à disposition. Par souci de simplification, nous appellerons prévision de l'activité, l'anticipation du poids et du nombre de lettres de transports. Dès lors, les prévisions de l'activité visent à fournir des éléments cruciaux pour :

- la planification des ressources matérielles et humaines,
- optimiser les règles de ramasse , d'expéditions et de distribution ,
- formaliser le comportement à court, moyen et long terme des clients,
- atteindre un niveau de service élevé,
- limiter la dépendance vis-à-vis de l'incertitude.

TFE étant un groupe et ayant comme souci d'homogénéiser les procédures de traitement, il convoite un système de prévision capable de s'adapter aux spécificités des différentes agences. Les prévisions doivent être facilement consultables, conviviales et accessibles par une interface web sur l'intranet. Enfin, les objectifs de la direction sont d'atteindre une erreur de prévision quotidienne inférieure à 5%.

Générer des prévisions à partir de modèles mathématiques ne se fait pas sans un historique. Début des années 2000, le groupe STEF-TFE investit massivement en moyens informatiques. Pour répondre aux exigences de traçabilités imposées par l'Europe (texte CE *n°* 178/2002), améliorer la rentabilité et du même coup le service rendu au client, le groupe utilise un Data Warehouse. C'est un entrepôt de données Oracle© dans lequel sont stockées toutes les informations relatives aux colis transportés et entreposés par le groupe. Les données proviennent d'un système d'information appelé GTI (Gestion Intégrée du Transport). Il regroupe les applications informatiques du groupe et alimente la base de données, soit plusieurs giga-octets d'information. L'information de référence est l'ordre de transport. Il donne accès aux informations de chargement et déchargement : produit, unité logistique, lieu de déchargement-chargement, date, heure, tiers expéditeur, tiers destinataire, nombre de colis, type d'emballage, poids, catégories d'emballages ... La base de données est mise à jour quotidiennement avec les données de la veille.

La suite de cet article propose une méthode de prévision qui combine méthodes endogènes et exogènes. La méthode cherche à anticiper, à un rythme quotidien et à un horizon de 15 jours, les valeurs de 3 chroniques dont la somme de deux d'entre elles est égale à la troisième. Pour retrouver une cohérence, les valeurs prévues indépendamment, sont redressées du haut vers le bas (« top-down aggregation »).

## 2 Le Modèle

### 2.1 Vue d'ensemble

Nous cherchons à modéliser trois séries temporelles. Elles reflètent les flux des marchandises sorties d'un quai. Les marchandises sont soit transportées vers un autre quai, c'est alors de l'expédition, soit livrées chez le destinataire final, c'est de la distribution. La somme des

## Prévisions d'activité à des fins opérationnelles

deux donne le total des denrées traitées à quai. Les deux quantités à prévoir, le nombre de lettres de transport et le poids des marchandises associées, sont relevés quotidiennement par TFE. La taille de l'historique est de cinq ans d'observations.

Le modèle statistique choisi cherche à se rapprocher le plus possible de l'activité du transporteur en la décomposant. Supposons que l'activité d'une agence TFE est déterminée par trois composantes, des facteurs déterministes (saisonnalité, jours fériés, promotions) dont les valeurs futures sont connues et des facteurs stochastiques. Ces dernières se décomposent en variables observées (grèves, perte ou gain d'un portefeuille client) dont les valeurs futures sont inconnues et en variables non observées (liquidation judiciaire du principal concurrent) dont les valeurs passées, présentes et à venir sont inconnues. Le modèle proposé tente d'extraire de la chronique les facteurs déterministes avant d'appliquer un modèle « autoprojectif » pour estimer les variables stochastiques observées. La différence entre les résultats obtenus par le modèle et les observations résulte des variables non observées ou mal estimées.

Soit le triplet  $(X_t, Y_t, Z_t) \in \mathbb{R}^3$  représente les valeurs à  $t$  des chroniques en « expédition », « distribution » et « total ». Elles sont soumises à la contrainte  $X + Y = Z$ . Ces valeurs sont calculées à partir d'autres chroniques  $(U_{X,t}, U_{Y,t}, U_{Z,t})$  par  $X_t = \omega_1 U_{X,t}$ ,  $Y_t = \omega_2 U_{Y,t}$  et  $Z_t = U_{Z,t}$ .  $\omega_1$  et  $\omega_2$  sont déterminés afin de satisfaire la contrainte.

$$U_{E,t}^1 = T_t S_t^1 (\beta^1 F)_t V_t^1 \epsilon_t^1 \quad (1)$$

$$U_{E,t}^2 = T_t S_t^2 (\beta^2 F)_t V_t^2 \epsilon_t^2 \quad (2)$$

avec  $E = \{X, Y, Z\}$ ,  $T$  la tendance,  $F$  un vecteur binaire correspondant à des événements calendaires,  $\beta$  leurs pondérations,  $S$  les coefficients saisonniers,  $V$  les processus stationnaires,  $\epsilon$  les bruits blancs. Pour étudier les composantes indépendamment les unes des autres, nous utilisons la fonction logarithme népérien.

$$\ln(U_{E,t}^1) = \ln(T_t) + \ln(S_t^1) + \ln((\beta^1 F)_t) + \ln(V_t^1) + \ln(\epsilon_t^1) \quad (3)$$

$$\ln(U_{E,t}^2) = \ln(T_t) + \ln(S_t^2) + \ln((\beta^2 F)_t) + \ln(V_t^2) + \ln(\epsilon_t^2) \quad (4)$$

Les résultats obtenus par les équations (3) et (4) sont combinée dans l'équation (5).

$$U_{E,t} = \lambda \ln(U_{E,t}^1) + (1 - \lambda) \ln(U_{E,t}^2) \quad (5)$$

## 2.2 Modélisation des éléments déterministes

### 2.2.1 La tendance

L'activité des agences TFE connaît deux tendances, une tendance intra-annuelle et une tendance inter-annuelle. La tendance intra-annuelle décrit l'activité d'une agence entre janvier et décembre. La tendance intra-annuelle est celle à long terme. Après les fortes dépenses de fin d'année, les ménages se remettent à économiser en janvier. C'est pourquoi l'activité est forte en décembre et chute en janvier. D'autre part, dans le milieu économique, les tendances sont lentes et progressives (Burtschy (1980)). Ces constations nous font choisir une la tendance linéaire par morceau de périodicité annuelle. Son équation s'écrit :

$$T_t = \theta t + \varphi An(t) + cste$$

avec  $An(t)$  l'année correspondante à la date  $t$ ;  $\theta t$  représente la tendance intra-annuelle,  $\varphi An(t)$  la tendance inter-annuelle.

### 2.2.2 La saisonnalité

L'activité des agences TFE est une superposition de mouvements oscillatoires de périodes hebdomadaires et journalières. Ainsi, elle admet une double saisonnalité que nous allons estimer pour l'effacer de la chronique. La saisonnalité hebdomadaire comporte 53 coefficients et la saisonnalité journalière en comporte 318 (53 semaines  $\times$  6 jours). La saisonnalité hebdomadaire est due à des périodes d'activités fluctuantes. Elles sont causées par des événements extérieurs tels que la météo, les vacances scolaires, les périodes de fête. Inclus dans la saisonnalité, il ne sera plus nécessaire de les analyser individuellement. L'idéal serait de s'affranchir du compteur des semaines pour ne retenir que les distances par rapport à des événements calendaires. Ainsi, nous n'aurions plus à nous préoccuper du fait que pendant l'année  $A$  tel férié est tombé la semaine  $s_{13}$  alors que l'année suivante il est tombé la semaine  $s_{14}$ . C'est une perspective à approfondir. Pour l'instant, nous retenons 53 semaines dont les deux extrêmes sont corrigés suivant le nombre de jours qu'elles comportent.

La saisonnalité journalière, très marquée, est due à une répartition de l'activité sur les 6 jours ouvrés de la semaine. Cette répartition dépend de l'agence en question. Le samedi par exemple, l'activité est réduite à son strict minimum, elle varie ensuite suivant les jours d'approvisionnements des grossistes et des GMS. Rappelons qu'ils représentent 80% des volumes transportés. Il existe de nombreuses méthodes de désaisonnalisation. Elles ont l'avantage de décrire l'activité aux décideurs. Les coefficients saisonniers montrent l'écart de la valeur moyenne constatée pour une semaine  $i$  et un jour  $j$  par rapport à la tendance. En matière de prévision des ventes, la méthode idéale n'existe pas. Partant du principe que deux valent mieux qu'une, pourquoi ne pas appliquer deux méthodes pour ne garder qu'une combinaison des résultats suivant le critère de minimisation de la variance des erreurs (voir paragraphe 2.4). La première méthode de dessaisonnalisation est celle des moyennes mobiles. La deuxième est une décomposition par régression linéaire.

La méthode des moyennes mobiles permet d'estimer des coefficients saisonniers selon les 3 étapes suivantes,

- calculer la série des moyennes mobiles centrées,
- calculer l'écart entre les valeurs observées et la moyenne mobile,
- normaliser les écarts, pour aboutir aux coefficients saisonniers.

Cette méthode est appliquée une première fois pour corriger les variations hebdomadaires et une deuxième fois pour corriger les variations journalières. Les résultats obtenus sont les coefficients saisonniers  $S^1$  de l'équation (1).

La deuxième méthode de décomposition est celle proposée par Buys-Ballot (1847). Elle consiste à trouver les coefficients  $S^2$  de l'équation (2) par MCO (Moindre Carrés Ordinaire).

$$\ln(U_t^2) - \ln(T_t) = \gamma_1 S_t^{2,1} + \gamma_2 S_t^{2,2} + \gamma_3 S_t^{2,3} + \gamma_4 S_t^{2,4} + \gamma_5 S_t^{2,5} + \gamma_6 S_t^{2,6} + \Phi_0 S_t^{2,0} + \dots + \Phi_{52} S_t^{2,52} + \zeta_t$$

La chronique diminuée de sa tendance, se décompose en une suite de composantes saisonnières correspondantes aux 6 jours de la semaine, aux 53 semaines de l'année et d'un processus

## Prévisions d'activité à des fins opérationnelles

$\zeta_t$ . Les  $p$  composantes saisonnières sont des variables binaires pour  $p$  saisons dans l'année. La variable binaire est égale à 1 lorsque la donnée se rapporte à la saison envisagée et 0 partout ailleurs.

### 2.2.3 Évènements calendaires

Les événements suivants sont appliqués sur les deux séries  $((\ln(U^i) - \ln(T) - \ln(S^i), i = \{1, 2\}))$ , corrigées des variations saisonnières et de la tendance, obtenues par moyennes mobiles et par la méthode de Buys-Ballot. La perte d'un jour d'activité provoque la récupération de cette activité sur les jours voisins. Par exemple, un jeudi férié peut conduire à une augmentation de l'activité le lundi par anticipation ou le vendredi par retard. Souvent, un jour férié a des conséquences prévisibles sur une période de 9 jours (J-4, J, J+4). Les conséquences sont différentes selon le jour férié, le jour de la semaine, l'agence en question.

La récupération de la perte d'un jour de travail sera variable selon que ce jour soit un lundi, un mardi, ou un autre. Si c'est un samedi, il y a peu d'activité à rattraper, alors que le lundi est une journée chargée. Si le férié tombe un vendredi, les GMS anticipent et demandent à être livrées le double jeudi. Les livraisons explosent le jeudi et les expéditions gonflent le mercredi pour des livraisons en A pour B. S'il tombe un lundi les GMS anticipent un peu sur le vendredi précédent et récupèrent surtout le mardi. Si le férié est un jeudi, il y a de fortes chances que l'activité du vendredi soit réduite, car les salariés font le pond. Le mercredi précédent sera d'autant plus chargé.

Un jour férié est souvent synonyme de fête ce qui engendre une augmentation de la consommation des ménages et par conséquent un renforcement de l'activité du transporteur. Mais cette hausse est variable selon qu'il s'agisse de la Toussaint ou de Noël. Les agences ne sont pas égales face à la hausse d'activité engendrée par un férié. L'agence de Bretanor par exemple, se trouve être située à côté d'une usine « Côte d'Or ». L'usine approvisionne tous les magasins de France en chocolat de Pâques. Cet approvisionnement commence des mois avant l'évènement et monopolise une grande partie des ressources de l'agence TFE. D'autres n'ont pas ce client et subissent moins de poussée ou elles la subissent à d'autres moments. Canal froid à Nantes par exemple transporte le muguet du 1er avril sur l'Hexagone.

Parfois, le 4<sup>ème</sup> jour suivant un férié est aussi le 2<sup>ème</sup> jour précédent un autre férié. C'est ce qui arrive en mai entre la fête du Travail (1er mai) et la Victoire 1945 (8 mai). Dans ce cas, il est difficile de séparer les effets issus des deux fériés.

Pour tenter de séparer ces 4 phénomènes engendrés par la tombée d'un jour férié nous retenons pour chacun d'eux les informations suivantes :

- nom du jour férié,
- jour de la semaine (lundi, ..., samedi),
- jour de la semaine des 4 jours précédents et des 4 jours suivants,
- éloignement des 8 jours encadrant le férié.

Il en résulte un tableau de 11 colonnes, une pour identifier le jour de la semaine, une codée entre -4 et 4 pour indiquer la distance du jour impacté au jour férié et les 9 autres pour reconnaître les fériés. Les  $T$  lignes représentent le nombre d'enregistrements de l'historique. Cette matrice est convertie en tableau disjonctif complet pour servir de variables binaires afin d'ajuster un modèle de régression sur la série désaisonnalisée :  $W = \beta^i F + \xi$ , avec  $i = \{1, 2\}$ ,  $W$  la chronique désaisonnalisée et  $F$  le tableau disjonctif complet. Cette pratique engendre un nombre important de variables (8 voisins  $\times$  9 jours fériés  $\times$  6 jours semaine). Toutes ne sont

pas significatives. Une comparaison entre la variance des estimations et la variance des erreurs (test de Fischer) permet de retenir les variables les plus discriminantes.

### 2.3 Modélisation des éléments stochastiques

Les procédures précédentes ont permis d'extraire de la chronique la tendance, les composantes saisonnières et les événements prévisibles. La série temporelle obtenue ( $V$ ) présente une relation de cause à effet entre l'observation à une date  $t$  et les observations précédentes ( $t-1$ ,  $t-2$ ,  $t-3$ ,  $t-4$ ,  $t-5$ ,  $t-6$ ). Pour modéliser et prévoir cette série, nous optons pour le lissage exponentiel simple qui a l'avantage d'être automatisable.  $\hat{V}_{T+1} = (1 - \alpha) \sum_{j=0}^{T-1} \alpha^j V_{T-j}$ . Il prend en compte les observations passées (de  $T-1$  à  $T-6$ ) et les pondère par la constante de lissage  $\alpha$ . Cette dernière est estimée de sorte à minimiser la différence carrée entre la chronique et les estimations du modèle.

Les valeurs prédites sont additionnées aux composantes déterministes que nous avons soustraites précédemment. Après application de la fonction exponentielle, nous obtenons l'estimation de la chronique d'origine.

### 2.4 Combinaison des prévisions

Du fait de deux procédures de décomposition, nous obtenons deux prévisions. L'erreur de prévision obtenue par les deux méthodes ne permet pas de conclure à la supériorité d'une sur l'autre. Le schéma de combinaison choisi, cherche à minimiser la variance de l'erreur de prévision résultant de la performance antérieure des prévisions individuelles (Bourbonnais et Usunier (2007)). Ne dépendant pas de la spécification d'un seul modèle, les prévisions combinées tentent à être plus robustes. La prévision combinée  $PC$ , est une moyenne pondérée des deux prévisions individuelles  $PU^1$  et  $PU^2$ ;  $PC = \lambda PU^1 + (1 - \lambda)PU^2$ ,  $\lambda$  est le coefficient pondérateur,  $0 < \lambda < 1$ . Soit  $EPC$ , l'erreur de prévision combinée,  $EPC = \lambda EPU^1 + (1 - \lambda)EPU^2$ , la variance de l'erreur de prévision est  $V(EPC) = \lambda^2 V(EPU^1) + (1 - \lambda)^2 V(EPU^2) + 2(1 - \lambda)\lambda COV(EPU^1, EPU^2)$ . Nous cherchons  $\lambda$  qui minimise  $V(EPC)$ , en annulant la dérivée première par rapport à  $\lambda$ . La solution est, si les erreurs de prévisions sont corrélées

$$\lambda = \frac{V(EPU^2) - COV(EPU^1, EPU^2)}{V(EPU^1) + V(EPU^2) - 2COV(EPU^1, EPU^2)}$$

,sinon

$$\lambda = \frac{V(EPU^2)}{V(EPU^1) + V(EPU^2)}$$

### 2.5 Redressement des prévisions et intervalle de prévision

Les chroniques du poids passé à quai en « expédition », « distribution » et au « total » sont prévues séparément. Or le total doit être égal à la somme des « expéditions » et « distributions ». Pour rendre cohérentes les prévisions, il faut redresser les chiffres. Par leur importance, les chiffres du « total » sont moins variants, donc plus aisés à prévoir. Partant du principe que l'erreur de prévision de la somme est moins élevée que les erreurs cumulées des « expéditions »



## Prévisions d'activité à des fins opérationnelles

et « distributions », nous choisissons de garder la prévision de la série « total des marchandises passées à quai » pour corriger les deux autres. Cette technique est appelée « top-down agrégation ». La procédure d'ajustement est la suivante, soit

- $\hat{X}_{t+h}$  le poids des expéditions prévues pour la date  $t + h$ ,
- $\hat{Y}_{t+h}$  le poids des distributions prévues pour la date  $t + h$ ,
- $\hat{Z}_{t+h}$  le poids total prévu pour la date  $t + h$ .

Nous cherchons  $\alpha$  et  $\beta$  tel que  $[\hat{Z}_{t+h} - (\alpha\hat{X}_{t+h} + \beta\hat{Y}_{t+h})]^2$  soit minimum, sous contrainte  $1,5 > \alpha > 0,5$  et  $1,5 > \beta > 0,5$ . Les contraintes indiquent que les prévisions en « expédition » ou « distribution » ne peuvent pas se tromper de plus de 50%. La méthode Quasi-Newton a été choisie pour résoudre le problème d'optimisation.

Pour accompagner les valeurs prévisionnelles, nous calculons les intervalles de prévision bilatéraux au niveau de probabilité de 95%. L'intervalle de prévision peut être utile lorsque qu'une décision est difficile à prendre. Il peut par exemple décider l'attribution d'un camion en plus sur une tournée. D'autre part, l'intervalle de prévision peut donner une idée de la confiance à accorder à la prévision. Plus l'intervalle s'écarte moins la prévision est stable.

## 3 Applications et résultats

Canal froid est une agence basée à Nantes. 60% de son activité se fait en distribution, contre 40% en expédition. Entre janvier et mars, l'activité est stable autour de 750 tonnes (du lundi au vendredi). Les jours fériés d'avril et mai, provoquent de brusques changements de régime. Le poids des marchandises transportées peut passer de moins de 10 tonnes un jour férié à plus de 1 110 tonnes deux jours avant. Les samedis et jours fériés exclus, l'activité d'avril et mai frôle les 1000 tonnes/jour. Juin est une période d'accalmie. Avec une activité autour de 900 tonnes/jours et quelques jours fériés, juillet et août sont des mois chargés. Septembre, octobre est une période qui retrouve le calme de début d'année avec une moyenne inférieure à 800 tonnes/jour. Enfin, l'activité progresse régulièrement tout le long de décembre pour atteindre un pic de 1 114 tonnes cinq jours avant Noël.

Le tableau 1 compare les résultats obtenus par le modèle avec les observations réelles. L'échantillon d'apprentissage est l'historique des poids précédant la date  $T$ .  $T$  varie entre le 01/01/07 et le 31/07/07 avec un pas de six jours. Nous prédisons les dates  $T + 1$  à  $T + 6$ . Nous comparons  $\hat{X}_{T+i}$ ,  $\hat{Y}_{T+i}$ ,  $\hat{Z}_{T+i}$  avec  $X_{T+i}$ ,  $Y_{T+i}$ ,  $Z_{T+i}$  pour  $i = \{1, \dots, 6\}$ .

Expédition		Distribution		Total	
BAR	DR	BAR	DR	BAR	DR
8%	10%	9%	12%	7%	9%

TAB. 1 – Indicateurs de confiance.

Les indicateurs d'évaluations sont, le biais absolu relatif (BAR) et la dispersion relative (DR). Les indicateurs se réfèrent à la moyenne des valeurs observées. Les résultats du modèle se trompent de 7% par rapport à la moyenne de la série « total des marchandises passées à

quai ». Les erreurs sont susceptibles de varier entre plus ou moins 9%. Notons que ces résultats sont meilleurs que ceux obtenus par les modèles existants chez TFE. Cependant, l'objectif d'une erreur de prévision annuelle moyenne, quelque soit l'agence, inférieure à 5%, n'est pas atteint. Mais on s'aperçoit que selon les agences TFE et les jours de la semaines, une prévisions juste à plus de 95% sur les 3 dernières semaines glissantes est réalisable.

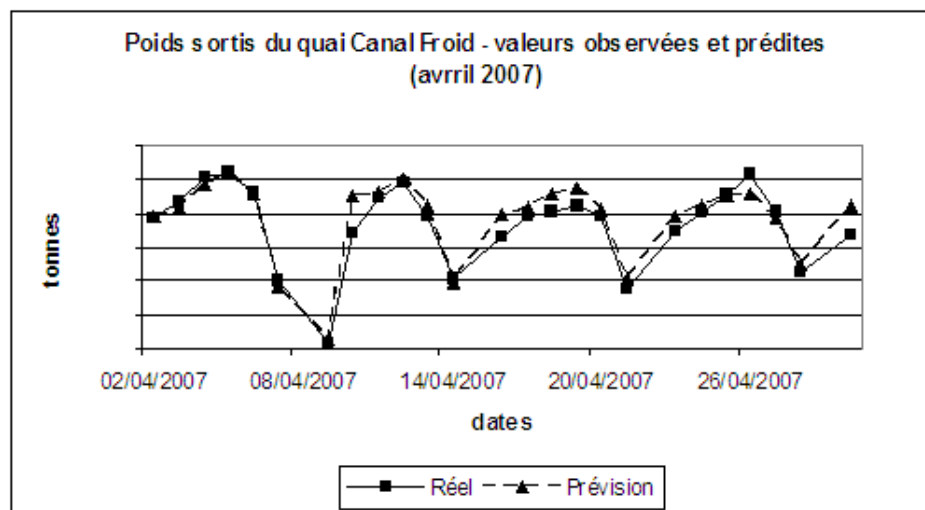


FIG. 1 – Comparaison observations réelles et prédites

Ajoutons que la qualité des prévisions varie suivant la bonne stabilité du comportement de l'activité. Une période d'activité stable (janvier à mars), permet d'extrapoler la chronique sans dégager d'erreurs importantes. Par contre, les périodes de turbulence (avril, mai) posent encore des problèmes (voir Fig. 1). Il apparaît que la méthode choisie pour mesurer l'effet des jours fériés sur l'activité n'est pas la plus adéquate. Pour être performante, la méthode nécessite plus d'historique. Rappelons qu'elle estime un coefficient de perte ou gain par MCO pour chacun des 4 jours passés et suivants un férié. Le cycle permettant de retrouver un même jour de la semaine pour un jour férié peut atteindre onze ans. D'autre part, les observations sont éloignées dans le temps et le comportement des ménages change au court du temps.

## 4 Conclusion et recherches futures

Les calculs de prévision sont centralisés au service d'informatique décisionnelle de STEF-TFE. Les résultats sont publiés sur une interface web spécialement créée pour ce besoin. Ils sont présentés sous forme de tableaux de bords et de graphiques. Les prévisions étant juste à plus de 90% en moyenne, elles sont une sérieuse aide à la décision.

Je suis conscient de la complexité du modèle dû au nombre de paramètres trop important au regard de la taille de l'historique (5 ans). En effet, pour chacune des décompositions il y a 2 paramètres pour la tendance, 318 pour la saisonnalité quotidienne, 53 pour l'hebdomadaire,

## Prévisions d'activité à des fins opérationnelles

20 pour les événements calendaires et un dernier pour le lissage exponentiel. L'avantage est de pouvoir expliquer au décideur l'effet de chacun des paramètres sur la quantité de marchandises sortie du quai. Une façon de réduire le nombre de paramètres est d'estimer la saisonnalité par une transformée de Fourier et de classer des événements calendaires en fonction de leurs effets. Il est également envisagé d'utiliser la corrélation qui peut exister entre les séries temporelles de 57 agences pour améliorer la modélisation. Les corrélations proviennent du fait que les agences TFE forment un réseau et s'expédient de la marchandise entre elles.

Une toute autre approche sera de désaisonnaliser par des filtres de Kalman emboîtés et introduire dans le modèle, sous forme d'impulsions, les effets des événements calendaires. C'est une méthode préconisée par Martin (1999) et appliquée sur la prévision de la consommation électrique.

Néanmoins, les économistes misent plus sur ce qu'ils appellent les « prévisions collaboratives ». Les statistiques ne suffisent pas pour obtenir une prévision fiable. Grâce au développement de la gestion partagée, il faut enrichir les résultats par la validation ou les commentaires des différents services de l'entreprise (la logistique, le marketing) et même des clients.

## Références

- Ayadi, S. (2005). *Le Supply Chain Management : Vers une optimisation globale des flux*. Working paper, Université Catholique de Lyon.
- Bourbonnais, R. et J. C. Usunier (2007). *Prévision des ventes, théorie et pratique*. Economica.
- Burtschy, B. e. M. C. (1980). A propos de prévision à court terme de la production industrielle. *Revue De Statistique Appliquée* tome 28 n° 2, 5–24.
- Buys-Ballot, C. H. D. (1847). Les changements périodiques de temperature.
- Martin, M.-M. (1999). Filtrage de kalman d'une série saisonnière, application à la prévision de la consommation d'électricité. *Revue De Statistique Appliquée* tome 47, n° 4, 69–86.
- Terrolle, C. (2004). *Évolution des rapports entre industriels et grande distribution : du partenariat à la satisfaction clients, vers l'émergence de nouvelles stratégies d'achat*. Mémoire de fin d'études, Université Paris I Pantéhon-Sorbonne.

## Summary

This article describes a operational research problem. A company that specializes in temperature controlled transportation wants optimize the planning of its human and material resources through short term activity forecasting. The challenge is to find a unique forecasting model adapted, without human intervention, to the specific needs of 57 company's officies. To do it, the company collected data since five years. To analyse them, mathematical algorithms for forecating time series are used. The work is to combine these tools to extract the maximum of determinist information that should be anticipated. The introduction presents the problem and its economic context. This is followed by a description of the process used and arguments to defend choices done. The adopted solutions are inventoried. Finally, the conclusion refers to courses of study.